International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

A bibliometric analysis of the model validation literature

Sibel Eker, Elena Rovenskaya, Simon Langan, Michael Obersteiner

International Institute for Applied Systems Analysis (IIASA)
Laxenburg, Austria

eker@iiasa.ac.at
@sibel_eker_

Model validation

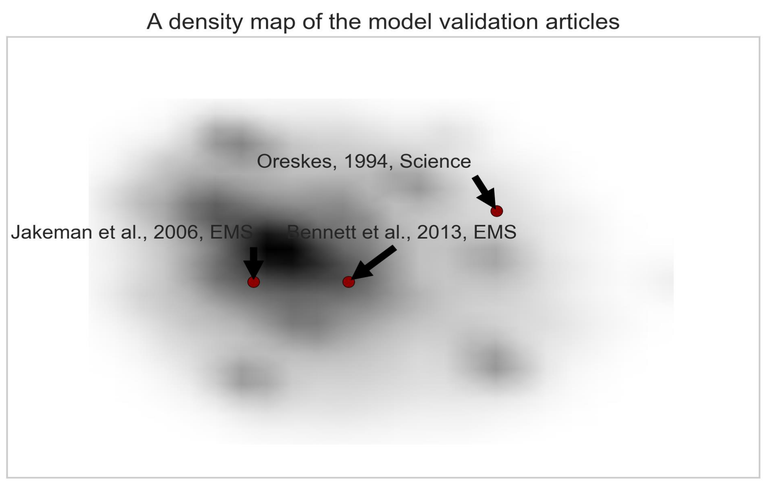
- Validation is a crucial step in quantitative modelling to establish confidence and reliability.
- However, it is often said that validation approaches proposed in the literature are not widely adopted by practitioners,
- and the validation approaches in different modeling fields do not benefit from each other.

This study…

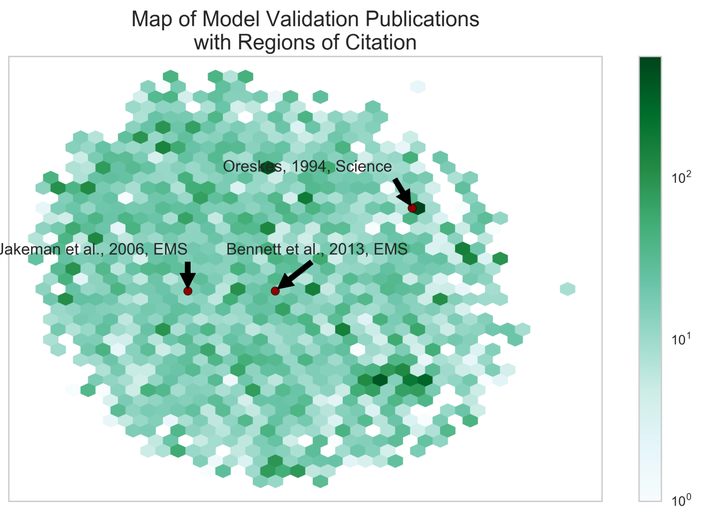
- examines to what extent validation (and specific validation approaches) is acknowledged and adopted,
- investigates how the validation practices in different modelling fields are related,
- employs a combination of citation and text-mining analyses on a dataset of 10688 academic publications.*

How related are these publications in terms of their content?

Based on nonlinear mapping** a large dense (dark) region contains many similar articles. Well-known articles are in the periphery, not very similar to the others, implying their content might not be widely adopted.



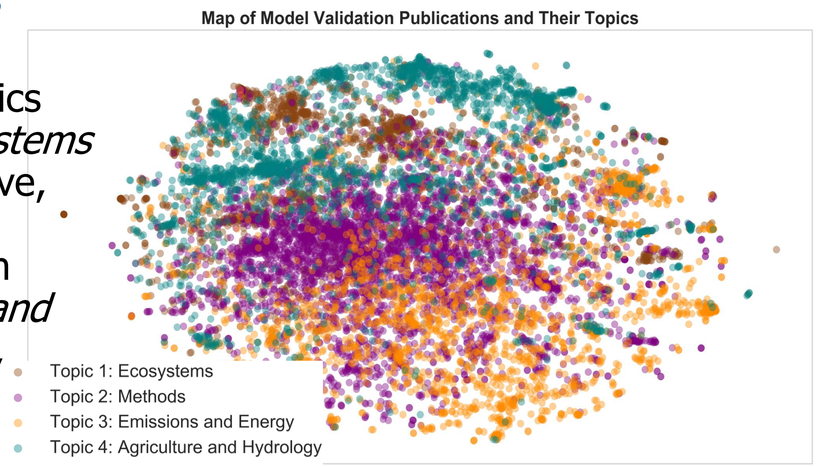
How does this relatedness reflect on citation scores as an indicator of uptake?



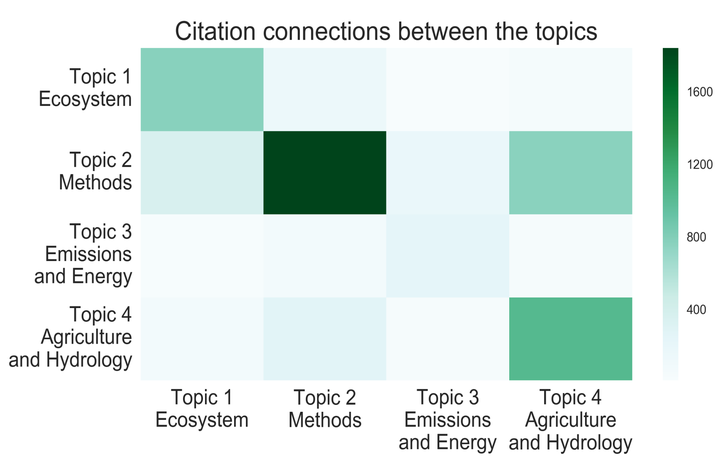
The most-cited articles are not necessarily in the densest regions. Instead, they are rather in the periphery of clusters, which can be considered different and more innovative.

Can this content relatedness be explained by different topics?

Among the four topics identified***, *Ecosystems* is relatively distinctive, implying a different content of validation articles. *Emissions and Energy* is dispersed, most similar to the *Methods* topic.

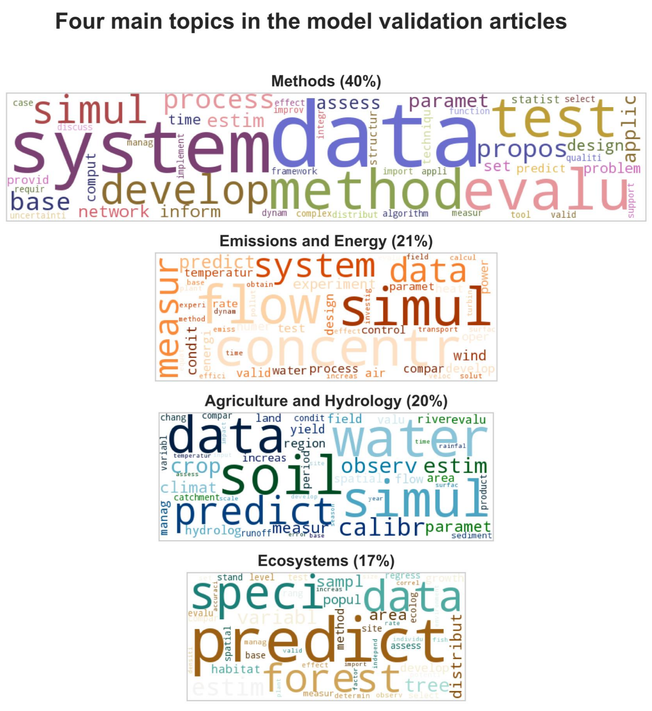


Do the publications from different fields cite each other?



The articles in each topic cite the articles in the same topic most, indicating that the validation literatures of these modeling areas are closed to each other. The *Methods* topic is the most-cited.

The word content of the four topics



METHODS

* The dataset is retrieved from Scopus with keywords "model validation, evaluation, assessment and testing" and limited to the disciplines such as environmental science, decision sciences, economics, energy, computer and social sciences.

** Using the t-SNE algorithm [4], which reduces the dimensions of multi-dimensional data points (articles) and builds a 2D map where the distances between the points depend on the word similarities in their abstract.

*** Topic modelling algorithm Latent Dirichlet Allocation (LDA) [5] allocates each document to one of the predefined number of bags to a certain extent, forming document-topic and topic-word pairs.

KEY FINDINGS

- Well-known articles proposing different validation approaches have a different content than most publications.
- The most-cited publications are not similar to the rest in terms of their content.
- Different modeling fields are closed to each other's validation practice.

References

1. Oreskes, N., Shrader-Frechette, K., Belitz, K., 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science* 263(5147) 641-646.Jakeman
2. Jakeman, A.J., Letcher, R.A., Norton, J.P., 2006. Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling & Software* 21(5) 602-614.
3. Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. *Environmental Modelling & Software* 40(Supplement C) 1-20.
4. Maaten, L.v.d., Hinton, G., 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research* 9(Nov) 2579-2605.
5. Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. *Journal of Machine Learning Research* 3(Jan) 993-1022.